Reprinted from:
I. Dodevski and A. Plückthun, Engineering Integral Membrane Proteins for Expression and Stability
In: Production of Membrane Proteins: Strategies for Expression and Isolation, Anne Skaja Robinson
Wiley-VCH Verlag GmbH & KGaA., 2011, pp. 277-296

10
Engineering Integral Membrane Proteins for Expression and Stability

Igor Dodevski and Andreas Plückthun

10.1
Introduction

Natural evolution provides integral membrane proteins (IMPs) with the necessary structural and biophysical properties to fulfill their function in the lipid bilayer of the cell. However, having evolved under the selective pressures of the cell and organism, IMPs have not been particularly adapted for high-level overproduction in laboratory expression hosts, let alone for high stability in detergent-solubilized form. These properties, as important as they appear in the daily work of the biochemist, have no relevance in the natural context of cellular function. These properties may be irrelevant for IMPs and they may even be selected against – some IMPs might need to be degraded and thus a limited stability might even be a desired natural trait.

To improve the process of producing a membrane protein at sufficient amounts and quality for structural studies, an ever-growing number of techniques for screening, selecting, and generating variations and mutations are being developed and implemented. The high pace of methodological developments reflects the observation that most wild-type IMPs are difficult to study structurally and biophysically – sometimes it is virtually impossible.

Currently, there are two different experimental strategies that are routinely used for increasing the chances of producing a well-expressed, soluble, and active membrane protein. The first strategy, which is the most intuitive – and probably part of every structure determination effort – relies on the extensive screening of a large number of experimental conditions for expression and solubilization of a given construct. Typically, many experimental parameters such as plasmid design, expression host, expression media, and temperatures, detergent, and buffer for solubilization have to be empirically optimized for any given target protein. These "classical" alterations may also include changes of fused tags, domains, or truncations and the use of a host of natural ligands binding to the protein. Nonetheless, it is frequently found that *none* of these conditions achieves the desired goal, as the protein itself is the limiting factor. In other words, for some proteins simply no conditions might exist in which this particular protein with its given sequence

Production of Membrane Proteins: Strategies for Expression and Isolation, First Edition.
Edited by Anne Skaja Robinson.

would show the required characteristics that would allow it to be studied in a purified, detergent-solubilized state.

The second strategy thus relies on changing the protein under study. Most popular has been the screening of protein homologs of a given target protein in order to identify a homolog that shows favorable biophysical properties. The homologs are usually screened in the context of a few globally optimized conditions for expression and solubilization. The most promising candidates for this strategy are proteins of prokaryotic origin, because the source of homologous sequences is large and diverse. Interestingly, many homolog screens tend to converge to sequences that derive from a small number of extremophile organisms, such as *Ralstonia metallidurans*, *Pyrococcus horikoshii*, or *Thermus thermophilus*. Clearly, extremophile organisms have evolved their protein repertoire to sustain extreme conditions (such as high temperature in the case of the thermophiles). Therefore, these proteins naturally show higher stability. The crystal structures of IMPs that derive from extremophile origin are over-represented among the set of solved atomic-resolution structures (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html). However, for many eukaryotic IMPs, including proteins of special medical relevance, such as the G-protein-coupled receptors (GPCRs), the strategy of homolog screening is not very useful, because there simply are no homologous sequences found in prokaryotic organisms.

Despite the strong research efforts put into these two strategies, they still often fail to establish appropriate conditions for expression and solubilization of a given IMP. Thus, the structural or biophysical characterization of many biologically relevant IMPs has remained a formidable challenge.

In such cases, the techniques of protein engineering can offer an alternative, third strategy. Protein engineering allows one to alter the biophysical properties of a given IMP itself, in order to increase its expression, solubility, stability, or crystal-forming propensity. It puts the focus on the protein sequence itself as an experimental parameter to be optimized. It is, in essence, a much "milder" approach than the search for homologs from thermophilic bacteria, but follows the same logic: protein engineering seeks to find a sequence *as close as possible* to the given starting sequence that bestows the desired properties on the IMP of interest.

By a process consisting of modifying the underlying amino acid sequence of a protein and screening or selecting by evolutionary strategies for improved protein versions, many roadblocks inherent to the production of IMPs can be removed. In this chapter we review the current experimental approaches for improving the biophysical properties of IMPs by protein engineering to produce better expressed and more stable IMPs.

10.2
Engineering Higher Expression

There are basically two ways of improving the biophysical properties of a given target protein by protein engineering. The first relies on modifying the amino acid

sequence of a protein by rational design. This approach usually requires a detailed structural description as well as a mechanistic understanding of the protein to be engineered. For many if not most IMPs, however, this information is not available. Almost by definition, a protein that is to be engineered in order to make it stable enough to have its structure determined will not have such information available.

However, even if a good structural model and a general mechanistic understanding of the protein was available, the sequence modifications required to rationally design better biophysical properties and higher stability are too complex to be predicted. Especially in IMPs, the detailed understanding of mutations that influence biosynthesis, folding, and aggregation is almost completely lacking.

The second protein engineering approach relies on a combination of random mutagenesis and selection or screening (see below for a definition of both terms and their distinction). This at first seemingly "irrational" approach is very powerful, provided sampling is very wide, and selection or screening is efficient and accurate. In other words, mutations tested must really include those that make the decisive difference, and selection or screening must be powerful enough to discern small improvements and to handle the commensurate number of mutants to be tested.

This strategy exploits the evolutionary principle of nature (i.e., random mutagenesis and selection, ideally in an iterative fashion) to generate a desired molecular property. This concept is known as directed protein evolution. Its strength lies in the fact that a desired molecular property can be obtained based on very low input information, provided the generation of diversity and the selection of the desired protein mutants is efficient. Since the iteration between diversification and selection allows one to explore the evolutionary potential of a protein in a combinatorial fashion, it is possible to evolve rather complex biophysical properties, such as the ones defining protein biosynthesis, membrane insertion, and folding in the membrane, usually summarized under the heading "expression."

The process of evolving such properties of a given IMP thus involves two experimental steps. In the first step, a diverse set of protein variants (a library) is generated by introducing mutations in the gene coding for the membrane protein of interest. Mutations can in principle comprise point mutations, insertions, or deletions. There are many standard molecular biology techniques by which such mutations can be introduced [1, 2] and those will not be reviewed here. Genetic diversity can be either concentrated to particular regions of the protein, to particular amino acid types [3] or the gene can be mutagenized randomly at a predefined error rate. Clearly, a strategy different from a random mutagenesis of the whole gene will require that additional information is available that justifies focusing of the efforts to particular regions.

In the second step, the library of mutant protein variants is analyzed with a functional assay in order to identify candidates showing the desired molecular property (e.g., higher expression and/or stability), termed "selection" or "screening." A technical distinction can be made between these two terms. *Screening* defines the analysis of individual mutants, which are kept separately, and whose

sequence identity is usually known before the experiment. As samples need to be handled in parallel, screening is restricted to cases where the number of mutant protein variants in the library is relatively small (in the range of hundreds). While several parameters can be determined in each sample, the challenges of sample handling puts a practical limit on sample numbers. Furthermore, individual samples of mutants will inevitably have some error in amounts, media, and concentrations, limiting the possibility of reliably identifying small differences.

If the sequence space that needs to be analyzed is much larger (e.g., in an attempt to evolve a particular protein), with no structural insight which regions to mutagenize preferentially, a screening of individually grown mutants is no more practical. With a library of millions of randomly mutagenized protein variants, it is simply not possible to screen each variant individually even with robotics equipment. In this case the methods of directed evolution, which are based on *selection* rather than screening, are more appropriate. The term "selection" defines methods in which all mutants are handled as a pool, in a single tube, and either a genetic selection (e.g., based on growth characteristics) or a physical selection (e.g., based on a direct measurement of receptor levels in an individual cell) is used to identify the clones with improved characteristics. In a *genetic* selection, the desired mutants have a growth advantage; in a *physical* selection, they are sorted or enriched by some physical separation.

Increasing the expression level of a target IMP represents a highly complex design task. We will not concentrate on promoter, transcription, or tags – even though they could in principle also be subjected to directed evolution – since they can usually be taken from other well-characterized working systems as a reasonable starting point. We will concentrate on the protein sequence itself. The protein sequence itself influences all the steps along the biosynthesis of an IMP: its translation rate (which might be influenced by the chosen codons), its incorporation into the membrane, its misfolding and aggregation in- and outside of the membrane during biosynthesis, as well as its susceptibility to aggregation or degradation after successful membrane insertion. All these steps influence the steady-state "expression level" of native protein. At the current state of knowledge, it is close to impossible to predict which mutations in a given amino acid sequence would influence any of these steps. We are therefore not able to engineer a molecular property called "functional expression level" by a rational, structure-based approach.

In the absence of such information it is necessary to include in the screening or selection as many mutants as possible to identify the ones that increase expression. In recent years two different approaches have been published that have proven to be very effective in evolving well-expressed IMPs starting from weakly expressed wild-type protein templates [4, 5].

10.2.1 Directed Evolution of a GPCR for Higher Expression

To directly address the importance of the protein sequence as an experimental parameter in IMP expression, we have developed a selection method in our labora-

tory based on the principles of directed evolution [5]. The method allows one to isolate well-expressed and functional GPCRs from libraries as big as 10^7–10^8 individual mutants, and it involves an iteration between diversification and selection. To demonstrate the power of the method we chose as the first example the neurotensin receptor NTR1, which expresses at about 800 functional receptors per *Escherichia coli* cell [6]. After several rounds of random mutagenesis and selection a receptor variant could be isolated that expresses about 10-fold more functional protein, and it also turned out to be more stable in detergents.

E.coli is the most convenient expression host for such experiments, since the transformation of large libraries is straightforward. As in the directed evolution approach, a diversified library will have to be brought into cells repeatedly; this ease of transformation is a very important consideration. Additionally, the handling of *E.coli* is very convenient on a large scale such that the final purification of the GPCR from *E.coli* is more easily implemented than with many other hosts.

The general method is schematically shown in Figure 10.1 for NTR1. (1) The wild-type NTR1 cDNA is randomly mutagenized by error-prone PCR (epPCR) to generate a library of NTR1 mutants. The mutant DNA is cloned into an expression vector, which previously had been optimized for the functional expression of NTR1 in *E.coli* [6]. The GPCR sequence is genetically fused to an N-terminal maltose-binding protein (MBP) and a C-terminal thioredoxin A (TrxA). The MBP, whose signal sequence will direct it to the Sec pathway, may enforce a periplasmic localization of the N-terminus of the GPCR, while the C-terminal TrxA may enforce a cytoplasmic localization of the C-terminus. (2) *E.coli* is transformed with the gene library and the proteins are expressed in the inner membrane of *E.coli* in liquid culture. Expression takes place at 20°C for 20h. (3) The cells are incubated with fluorescently labeled ligand BODIPY-neurotensin (BODIPY-NT), which shows high specificity and affinity for the neurotensin receptor. To allow binding of the fluorescent ligand to the receptors expressed in the inner membrane of *E.coli*, the outer membrane has to be partially permeabilized by an appropriate permeabilization buffer. (4) The cells expressing the highest number of functional receptors, which therefore exhibit the greatest fluorescence, are sorted by fluorescence activated cell sorting (FACS). The cells are sorted directly into growth medium and can directly be cultivated for a next round of sorting of the highest expressing mutants. By repeating the cycle of receptor expression and sorting, the highest expressing mutants can be strongly enriched from the large pool of initial mutants. Since all mutants are in the same test tube, any small variation in permeabilization efficiency of the buffer or other variations in concentrations and times will affect all mutants equally. Therefore, this approach can be thought of as a competitive experiment between different cells, amplifying small differences by serial repetition.

Whenever additional genetic diversity is desired after any FACS round, the plasmid DNA of the selected mutants is isolated and the GPCR sequence is further randomized by epPCR. The FACS selection is then repeated.

The procedure for evolving the expression level of NTR1 is outlined in Figure 10.2. The initial randomized NTR1 library was subjected to four rounds of FACS.

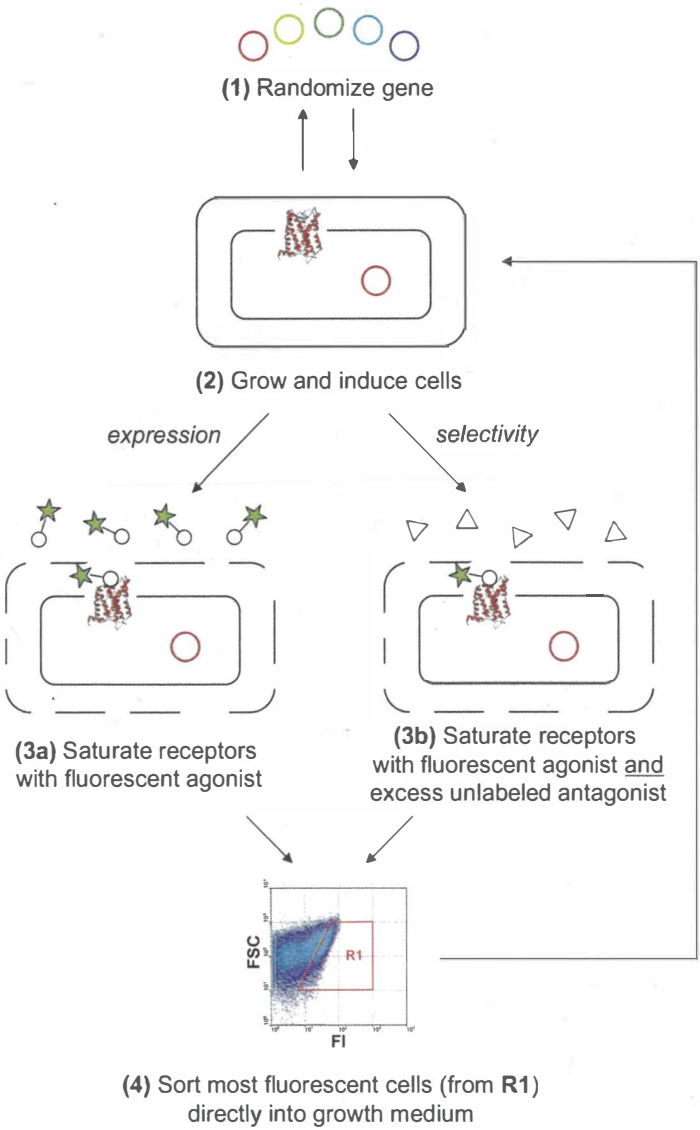

Figure 10.1 General selection scheme for increasing expression level (steps 1, 2, 3a, 4, back to 2) or altering ligand selectivity (steps 1, 2, 3b, 4, back to 2).

In each round, only the most fluorescent 0.1–1% of the cells were collected. Nonetheless, after these rounds, the evolved pool had a mean fluorescence intensity (MFI) no greater than that of the wild-type sequence. epPCR was used to overlay another set of random mutations on top of those that were enriched after the first four rounds of FACS and this rerandomized library was again subjected to four rounds of sorting. In this second set of sorts, the MFI of the pool overtook that of

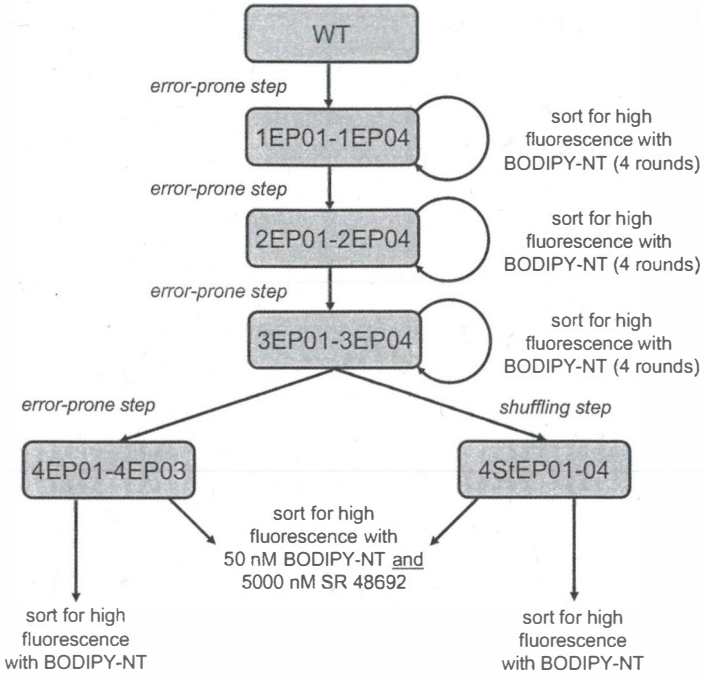

Figure 10.2 Flowchart for selections of NTR1 variants leading to increased expression level or altered ligand selectivity. WT = wild-type.

wild-type NTR1. After a third randomization step followed by four more rounds of FACS, the evolved pool was split into two. One half was randomized by epPCR a fourth time and the other half was subjected to DNA shuffling, using the staggered extension process [7]. After these selections, the MFI was approximately 5 times that of wild-type NTR1.

From the enriched "error-prone" pool (4EP03 pool), 48 single clones were analyzed for receptor expression level. Figure 10.3 shows that the clone with the best functional receptor expression level per cell, D03, exhibited approximately a 10-fold increase in specific signal, as assayed by [^{3}H]neurotensin binding. This shows that the receptor has not "adapted" to the fluorescent dye, as binding was assessed with unmodified ligand. The receptor shows nine amino acid mutations compared to wild-type NTR1. Analysis of the expression level of D03 in eukaryotic expression hosts shows that it also expressed about 12-fold better in *Pichia pastoris* and 3-fold better in HEK293T cells. These increased expression levels show that the receptor has also not "adapted" to a prokaryotic expression host. With respect to protein purification, D03 can functionally and quantitatively be solubilized in detergent micelles from the inner membrane of *E. coli* and yields about 0.5 mg of solubilized and functional GPCR per liter of shaking flask expression culture (around 4–5 g cells wet weight). Therefore, the increased functional level seen in whole cells is maintained after solubilization and purification.

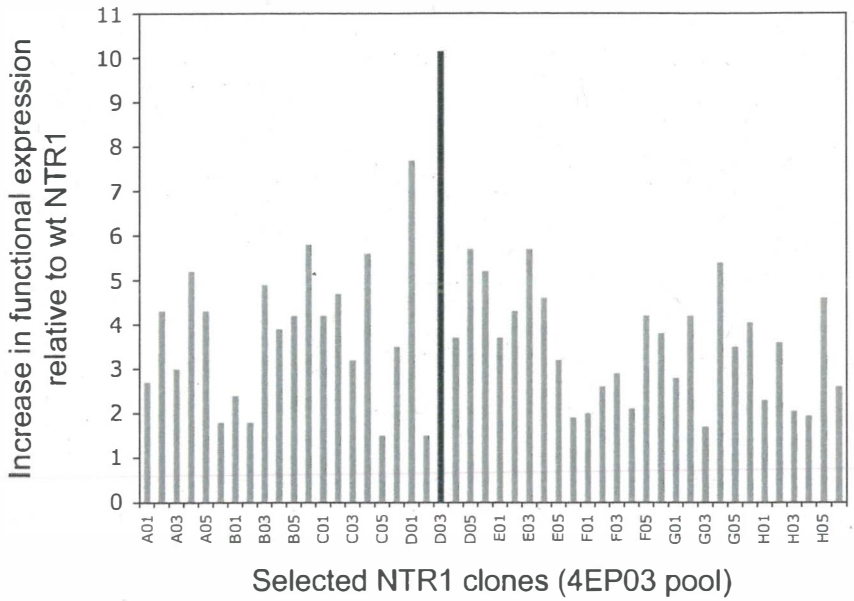

Figure 10.3 Radioligand binding analysis of 48 clones isolated from selections for maximum receptor expression level (4EP03 pool, four rounds of epPCR, each followed by three to four rounds of sorting by FACS). The bars represent receptor expression level per cell, relative to wild-type NTR1 expression level. D03 (black bar) is the clone showing the highest expression level (10-fold that of wild-type NTR1).

The evolved receptor D03 also retains the biochemical and pharmacological properties of wild-type NTR1. As D03 had been evolved under selective pressure to conserve a functional ligand-binding site, it retains the same affinity for the high-affinity agonist neurotensin as wild-type NTR1 (K_d ~ 0.1 nM in intact *E.coli* cells). Note that selection also has not improved the affinity. This can be explained: since the fluorescent dye was present at 50 nM during the selection, the receptor should be essentially saturated such that higher affinity mutants would not be rewarded. Neurotensin binding can also be competed by a molar excess of the antagonist SR48692, indicating that the ligand-binding site has indeed been conserved during the selections.

In addition to the selections performed for increasing the expression level, the selection method has also been applied to altering the binding selectivity of wild-type NTR1 (Figure 10.2). A motivation for such an experiment can be to favor one conformation of the GPCR over others. These selections were done by incubating a mutant library with a 100-fold excess of unlabeled antagonist SR48692, in addition to the labeled agonist BODIPY-NT, and selecting the cells showing the highest fluorescence. Those receptor mutants are thus expected to still bind the labeled agonist, but not the antagonist competitor. The selected mutants showed one strong consensus mutation in TM7 (F358S). The effects of mutation at Phe358

had been studied by others [8, 9] and reveal that a substitution to alanine at this position results not only in decreased antagonist affinity, but also in spontaneous basal inositol phosphate production in a receptor-dependent manner. Thus, more generally, the selection method by FACS has the ability to rapidly isolate mutations that may trap receptors in the active or inactive state. This approach is thus complementary to the engineering of GPCR stability, reviewed in Section 10.3.2.

There are several technical advantages inherent to this FACS-based selection method. Most importantly, for the first time IMPs can be evolved under laboratory conditions by harnessing the full power of the natural selection process. By physically linking the genotype to the phenotype – by containing the gene encoding the corresponding receptor in an *E. coli* cell – it is possible to move from a low-throughput *screening* set-up to a high-throughput *selection* setup. This allows one to sample a much larger sequence space in the range of 10^7–10^8 individual mutants per selection round as opposed to only thousands of mutants, when they have to be screened one-by-one. Being able to sample the largest sequence space possible strongly increases the probability of finding rare expression enhancing mutations. Moreover, since directed evolution is an iterative process, selection is clearly advantageous over screening. The selection from a large pool has to be done after each library diversification – a prohibitive amount of work when screening single mutants.

A second advantage, which makes the application of the selection method very appealing, is that there is a strong selection pressure for correctly folded, functional receptors inherent in the selection for the number of active receptors. Well-expressing receptor variants are selected only if the ligand-binding site remains conserved throughout the evolutionary process, because in order to be selected the receptor variants must bind the fluorescent ligand. Receptor variants adopting non-native receptor folds are selected against, even if they are well expressed. For GPCRs, ligand binding is a very strong indicator of a correctly folded protein. The ligand-binding sites of GPCRs recognizing small ligands are mostly contributed by several amino acids located on different transmembrane helices. These different helices must therefore be in a wild-type-like conformation to provide a functional high-affinity ligand-binding site.

The third advantage in implementing this method relies on the fact that the expression level of a given IMP can be improved without having to make assumptions about the mechanism of the expression process or about the amino acid substitutions that might influence it, or even the regions in the sequence that would affect it. The combinatorial approach inherent in the method allows evolving the desired property in a nearly assumption-free manner. This is a very important consideration, as many sequence correlations probably have yet to be discovered and others may be incomplete or even incorrect.

The method's potential to select for functional receptors by using fluorescent ligands is – at the moment – compromised by the fact that fluorescent ligands are not yet available for every GPCR. To date, the chemical synthesis of fluorescent ligands has been described in the literature for maybe 30 GPCRs or so [10, 11]. Fortunately, the ever-increasing research activity in the field of GPCRs and IMP

biochemistry in general makes it very likely that new fluorescently labeled ligands will be developed as powerful tools in membrane protein research. Moreover, many GPCR ligands do not require complicated chemistry for their derivatization with fluorophores. Some can be synthesized in any laboratory that is equipped with a basic equipment for chemical synthesis; this is especially true for peptide ligands, which provide a variety of possibilities for chemical derivatization. The N-terminal amino groups, ε-amino groups of lysine residues, thiol groups of cysteine residues, or the C-terminal carboxyl groups can often be coupled to fluorophores without considerably compromising the binding affinity of the peptide ligand.

While the selection method by FACS is able to evolve well-expressed GPCRs that conserve their functional ligand-binding site (a very strong criterion for a correctly folded GPCR), other functional characteristics of these proteins are more difficult to enforce to be conserved during the directed evolution process. Most importantly, the ability to transmit binding signals (i.e., to couple to G-proteins) is very difficult to select as a property to be retained, as the selection process takes place in *E. coli*. Nevertheless, in the example of the NTR1 receptor studied, the ability to signal via G-protein *was* retained [5]. This was determined experimentally by recloning the selected mutants to mammalian cells and measure the ability to signal via G-proteins directly. Thus, it is important to characterize evolved mutants very carefully after the selection experiment, as is generally the case for every protein engineering method presented in this chapter.

To prove the usefulness of evolved mutants for structural studies their phenotypes should be carefully compared to their wild-type progenitor. When we performed this analysis for D03, we found that the signaling capability was retained, but somewhat more agonist was needed, compared to wild-type NTR1. This was found in mammalian cells by measuring Ca^{2+} mobilization via coupling to $G_{q/11}$. When one of the evolved mutations in D03 (R167L in the conserved (D/E)R(W/Y) motif in TM3) was reverted to wild-type Arg167, D03 would show a similar signaling behavior as wild-type NTR1 [5]. This analysis underlines the importance of performing a careful functional characterization of evolved mutants. Nonetheless, it also emphasizes that the basic functions, such as agonist binding, antagonist binding, and signaling via the G-proteins, can be maintained in these evolved GPCRs.

10.2.2
Increasing Expression by Random Mutagenesis and Dot-Blot Based Screening

A second screening method for increasing expression levels of IMPs has been developed by the Nordlund lab [4]. The colony filtration (CoFi) blot method is well suited to screen somewhat smaller mutant libraries for expression, since thousands of single mutants can be screened simultaneously. However, since single colonies must be discernable on plates, it is difficult to extend this to very large libraries. By applying the CoFi blot method to several membrane proteins their expression level could be increased. After one round of random mutagenesis and

screening, mutants could be identified that express 1.5- to 40-fold better than the corresponding wild-type protein, depending on the protein under study and the initial expression level.

The CoFi blot has originally been developed for screening the expression level of soluble proteins in *E. coli* [12]. It has proven very efficient in identifying mutants of eukaryotic proteins that show higher soluble expression yields than the wild-type proteins. Recently, the method has been adapted to allow screening of mutant libraries of membrane proteins [4]. The strategy consists of four steps:

i) A library of mutants is generated by random mutagenesis of the wild-type open reading frame used as template DNA, cloning the library into an expression vector (containing an N-terminal FLAG tag and a C-terminal His_6 tag), and transforming *E. coli* cells. The cells are plated onto an LB-agar plate and grown until medium-sized colonies are visible.

ii) The colonies are picked up onto a Durapore filter membrane by overlaying the colonies with the membrane and peeling them off the agar, as they will stick to the membrane. Expression is then induced by placing the membrane onto an LB-agar plate containing isopropyl-β-D-thiogalactopyranoside.

iii) The membrane is placed on top of a nitrocellulose membrane and a Whatman 3MM paper soaked in lysis buffer containing detergents. Solubilized membrane proteins will diffuse through the filter membrane (Durapore) and are captured on the nitrocellulose membrane.

iv) The amount of expressed and solubilized membrane protein is quantified by probing the nitrocellulose membrane with a His-tag-specific reagent using standard equipment for Western blotting. The mutant colonies' expression level can easily be compared to a wild-type reference and the best-expressing mutants can be identified.

The method was benchmarked by subjecting nine membrane proteins to one round of random mutagenesis and screening the mutant libraries for expression. The set of target proteins consisted of eight prokaryotic proteins and one human protein, and they were from different functional classes. They were classified as showing either no, low-, or medium-level expression. For five of the nine target proteins (all classified as low or medium expressing proteins) the method was able to identify better expressing mutants. For an *E. coli* glycosyl transferase the expression yield was improved by an impressive 40-fold over the wild-type level (from about 25 µg/l to 1 mg/l of purified protein) as a result of three amino acid mutations. For the human microsomal glutathione S-transferase 2, several single amino acid mutations could be identified that increase the expression level 2-fold. These results are very encouraging because they are based on only one round of mutagenesis and screening.

The ability to apply the method to nine different proteins from different functional classes in a benchmark test exemplifies the biggest advantage of the CoFi blot method–its generality. In principle, it can be applied to any membrane

protein with no prior structural or functional information about the target protein. Moreover, the method is easy to implement and the costs are low. However, the biggest strength of the method–its simplicity and generality–represents at the same time its greatest potential weakness. Random mutagenesis of a target gene is expected to generate mostly misfolded and nonfunctional protein mutants. As the C-terminal His_6 tag is used as the sole indicator of expression, the structural and functional state of the protein mutants is completely neglected in the screen. Reducing the screening criteria to His_6 tag detection entails a considerable risk of identifying well-expressed, but misfolded proteins. Since there is no functional screening criterion directly implemented into the method, it is crucial to carefully characterize the functional state of improved mutants after the screen. With regard to this risk, the developers of the CoFi blot analyzed the catalytic activity of selected mutants of one of their target proteins–the human microsomal glutathione S-transferase 2. The two analyzed mutants both showed only one amino acid mutation and their catalytic activities were comparable to the wild-type protein. Despite the encouraging results on these single-amino-acid mutants of one membrane protein, the risk of losing functionality, when proteins are evolved in the absence of selective pressure for functionality, remains considerably high. It is very likely that this risk increases when multiple mutations start to accumulate in experiments in which several rounds of mutagenesis and selection are performed.

10.3 Engineering Higher Stability

Establishing high expression levels of correctly folded IMPs is only the first critical step in the process of producing sufficient amounts of functional protein for biophysical and structural studies. Equally critical is the protein purification process as well as the stability of the purified protein in the solubilized state, for example, for structural studies, where the detergent-solubilized protein will be studied at high concentrations for extended times. Even though stability and functional expression yield show some correlation (Schlinkmann *et al.*, unpublished; Dodevski *et al.*, unpublished), the properties are not identical and, thus, stability must be tested separately.

Unlike soluble globular proteins, the purification process of IMPs and most methods of biophysical analysis absolutely require the help of detergent molecules to extract them from their natural physical environment–the phospholipid membrane–and to transfer them into an isotropic solution, established by detergent micelles. This process is problematic from a thermodynamic point of view because the physical environment provided by the detergent micelle is very different from the phospholipid membrane. Membrane protein solubilization by detergents therefore frequently leads to protein unfolding, aggregation, and loss of function.

For GPCRs, for example, the solubilization process turns out to be particularly destructive because of their marginal biophysical stability. GPCRs naturally exist

as structurally flexible molecules, as conformational changes are required to exert their function in transducing extracellular signals across the phospholipid bilayer. The observation that GPCRs can activate G-proteins even in the absence of an activating agonist ligand – a phenomenon called basal activity – implies that the receptors can continuously adopt different receptor conformations that are of similar thermodynamic energy and separated by rather small energetic barriers. This structural and thermodynamic heterogeneity is most likely responsible for the rapid receptor unfolding observed upon receptor solubilization with detergents. The marginal stability of GPCRs in detergent micelles represents the rule for IMPs, rather than the exception. In fact, the activity of most IMPs is linked to conformational flexibility in certain regions of the proteins. Most IMPs tend to quickly lose their activity when solubilized in detergent micelles, because of limited biophysical stability.

The protein engineer's contribution to solve this problem is to identify stability-enhancing mutations in the amino acid sequence of a target protein. Here, we define "stability" as the molecular property of the protein to maintain a correctly folded and active conformation in detergent-solubilized form. We therefore set out to identify mutations that increase the half-life of a protein in detergent micelles. At the current state of knowledge and in the absence of high-resolution structural information for the great majority of IMPs, it is close to impossible to predict in advance what mutations will improve the stability of a given protein target. Stated more practically, the identification of stability-enhancing mutants relies on setting up appropriate screening experiments for rapidly and reliably analyzing collections of many mutants.

10.3.1
Stabilizing a Prokaryotic IMP by Cysteine-Scanning, Random Mutagenesis, and Screening in a 96-Well Assay Format

The approach of using protein engineering for stabilizing IMPs in detergent micelles gained momentum with the findings by the Bowie lab published in 1999, namely that stability-enhancing mutations were not rare at all in IMPs [13]. In the analysis of 20 single cysteine-substituted mutants of the *E. coli* diacylglycerol kinase (DGK), two mutants showed significantly higher resistance to thermal inactivation. Combining the two mutations in a single protein revealed a partly additive effect and the stability was further increased. Strikingly, while the half-life of the double mutant at 70 °C is 51 min, the half-life of wild-type DGK is less than 1 min. In a follow-up study on DGK, a collection of 1560 random mutants was screened for thermal stability [14]. Twelve different single mutants of DGK showed higher stability in detergent solution. The four most stabilizing mutations were combined to construct the quadruple mutant CLLD-DGK, which showed a half-life of 35 min at 80 °C. This is about 18 times that of the most stable single DGK mutant. Most importantly, the stabilized mutant showed similar catalytic activity as the wild-type.

How were the 1560 single clones screened to identify stability-enhancing mutations? The screening setup is relatively simple. The mutants were expressed in

96-deep-well plates and the protein was solubilized with detergents. In a 96-well format an aliquot was directly assayed for DGK enzymatic activity in a colorimetric assay. A second aliquot was assayed after it had been exposed to heat for a defined time period to inactivate the protein. Mutants showing higher activity than wild-type after the inactivation step were selected and further characterized. The finding that stability-enhancing mutations in IMPs may be identified relatively easily has inspired other research groups, including our own, to explore the potential of protein engineering for improving the biophysical properties of IMPs, and to devise additional methods to screen for stability for IMPs without enzymatic activity. In this respect, GPCRs have been of special interest.

10.3.2
Stabilizing GPCRs by Alanine-Scanning and Single-Clone Screening

The Tate lab has applied an alanine-scan to identify mutants of GPCRs showing increased thermal stability in detergent solution. This strategy has successfully been implemented for three different GPCRs [15–17] – a class of membrane proteins that is notoriously unstable in detergent solution. The most significant example of their work is the engineering of a β_1-adrenergic receptor mutant, which is highly stable in the detergent octylglucoside. The combination of alanine mutations rendered the receptor stable enough to be crystallized and to determine its atomic resolution structure [18].

Their strategy for identifying stability-enhancing mutations in a GPCR is based on a relatively simple methodology. A collection of mutants was prepared by mutating each position in a receptor to an alanine residue. If the wild-type residue was already alanine, leucine was introduced, as this is a helix-forming residue compatible with a membrane location. The single-point mutants were then analyzed for their potential to increase thermal stability of the receptor in detergent solution. The thermal stability of individual receptor mutants was determined on unpurified samples after detergent solubilization of whole cells. To measure the stability of each mutant, one receptor aliquot was heated for a fixed period of time, a second aliquot was kept on ice. Both of the samples were then assayed for their content of folded receptor by a radioligand binding assay (LBA). The best single mutants were then combined to test for additivity and to produce more stabilized receptor variants.

The alanine-screening methodology was successfully applied to three different GPCRs. In the case of the adenosine receptor A_{2A}, two different thermostabilized mutants were constructed [17]. For the mutant A2a-rant21, the melting temperature (T_m) was 17 °C higher than that of the wild-type, for the mutant A2a-rag23 the improvement was 9 °C. Each receptor shows four mutations. Interestingly, the two mutants seem to be stabilized in different conformations: A2a-rant21 preferentially adopts an antagonist-binding state, A2a-rag23 prefers an agonist-binding state. As a second receptor, NTR1 was stabilized as a result of four mutations [15]. The NTR1 mutant termed "NTS1-7m" was 17 °C more stable than wild-type NTR1. Lastly, the turkey β_1-adrenergic receptor was stabilized by combining six mutations

[16]. The mutant b_1AR-m23 was 23 °C more stable than the wild-type protein. This mutant was then crystallized in octylglucoside and its atomic resolution structure was determined in the presence of the antagonist cyanopindolol.

The high success rate of identifying stabilizing mutations in both scanning mutagenesis strategies, with alanine scanning or cysteine scanning, are encouraging in that stabilizing mutations appear to be very frequent. However, both strategies suffer form a basic limitation. Their application is limited to those membrane proteins that show a relatively high wild-type expression level. This excludes many membrane proteins that fail to be expressed above a certain threshold level. Furthermore, the linear scanning against one type of amino acid will only detect positions where an unfavorable amino acid needs to be removed. Since a full screen is not carried out against all substitutions, amino acids able to make a new interaction, or those filling out a cavity better, would not be discovered. Since these methods are screening methods where each mutant is expressed and solubilized separately, it would hardly be feasible to increase the throughput to the required scale.

10.3.3
Stabilizing GPCRs by Random Mutagenesis and Screening in a 96-Well Assay Format

In Section 10.2.1 we described a high-throughput selection method with the main goal of identifying *expression*-enhancing mutations in GPCRs. An advantage of this strategy is that it generates large sets of well-expressing and functional mutants. We wondered if there was a quick and reliable way of screening these collections of mutants for their potential of increasing thermal stability in addition to increasing expression. As we wanted to perform the stability screen on as many receptor mutants as possible, we had to revise the LBA, which is the rate-limiting step of the conventional stability screening method for GPCRs (which had conventionally been employed in the alanine-scanning method by Tate [15–17] and previously by us [5]). In the conventional LBA method, each sample has to be processed by a small size-exclusion column (typically a spin column) to separate bound from unbound ligand and assess the ligand-binding signal, since solubilized receptor cannot be quantitatively bound to filters. This spin-column step cannot be performed in a 96-well assay format and therefore strongly limits the assay throughput. The key feature of the newly developed method [19] presented below is the immobilization of biotinylated receptor on streptavidin-coated paramagnetic beads. By immobilizing the receptor, all essential experimental steps of the stability screen–purification, exposure to heat, and LBA–can be performed with small receptor amounts and in a highly parallelized 96-well format. Immobilized receptor can easily be separated from detergent-solubilized lysates of whole cells by magnetic force, which yields highly concentrated and purified receptor preparations. Most importantly, magnetic capturing also allows for a convenient separation of bound from unbound ligand in the LBA, which avoids the handling of size exclusion spin columns. All essential steps can therefore be performed in a 96-well assay format [19].

The method for screening thermal stability of GPCRs in a 96-well assay format [19] consists of four steps. (i) The receptor mutants are expressed in *E.coli* and biotinylated *in vivo*. (ii) The receptors are solubilized and partially purified by immobilization on streptavidin-coated paramagnetic beads. (iii) The receptors are exposed to stability screening conditions that induce receptor unfolding (e.g., heat, detergent, buffer). (iv) The amount of residual folded receptor is determined by a LBA after exposing the receptor to the stability screening conditions. Comparison of the amount of correctly folded receptor before and after heat treatment yields a stability index for each mutant. The stability index is calculated by dividing the residual amount of receptor that has been exposed to the harsh conditions (e.g., high temperature or specific detergent) by the initial amount that is determined from an aliquot of nonexposed beads.

We applied this method to 96 randomly picked clones of a library of NTR1 mutants, which had been evolved for higher expression by FACS [5]. The mutants had gone through four rounds of random mutagenesis and after each round they had been sorted for highest expression by FACS for three to four rounds. The selected mutants showed on average a 5-fold higher functional expression level than wild-type NTR1 as a result of an average of nine amino acid mutations per receptor. For screening their stability, the mutants were expressed in 24-well plates, solubilized in the presence of a detergent mixture containing dodecyl maltoside (DDM), 3-(3-cholamidopropyl)-dimethylammoniopropane sulfonate (CHAPS), and cholesteryl hemisuccinate (CHS), and immobilized on streptavidin-coated magnetic beads in a 96-well plate. After washing the beads with the help of magnetic capturing, equal amounts of beads of each purified mutant were dispensed into two 96-well thermocycler plates. One plate was kept on ice. The second plate was placed in a thermocycler and exposed to 37 °C for 20 min to induce receptor inactivation. After cooling the heated plate on ice, a radioligand binding assay was performed on both plates to determine the amount of correctly folded receptor for each mutant before and after the heat inactivation step. After incubation with radioligand, unbound radioligand can be separated from bound ligand by magnetic capturing of the beads. A major advantage of this new screening method is that the LBA is performed directly in the wells of a thermocycler plate. For measuring the radioactivity the beads are simply transferred to a 96-well plate containing liquid scintillation cocktail.

Two screens were performed on the 96 mutants to identify the most stable mutants in either of the detergents, DDM or decyl maltoside (DM). Sixteen mutants showed higher thermostability than wild-type NTR1 in both detergents. Figure 10.4 shows that the difference in stability between the mutants and wild-type NTR1 was more pronounced in DM (Figure 10.4b) than in DDM (Figure 10.4a). The two clones that performed best in the screens (clones 70 and 73) were further characterized to get a better picture of the stability improvement. Melting curves in Figure 10.4(c) were recorded by exposing the mutants in DM, CHAPS, and CHS to different temperatures for 20 min. For both mutants the melting temperature T_m was increased by 6 °C compared to wild-type NTR1. The relatively moderate increase in stability most probably reflects the fact that the original pool

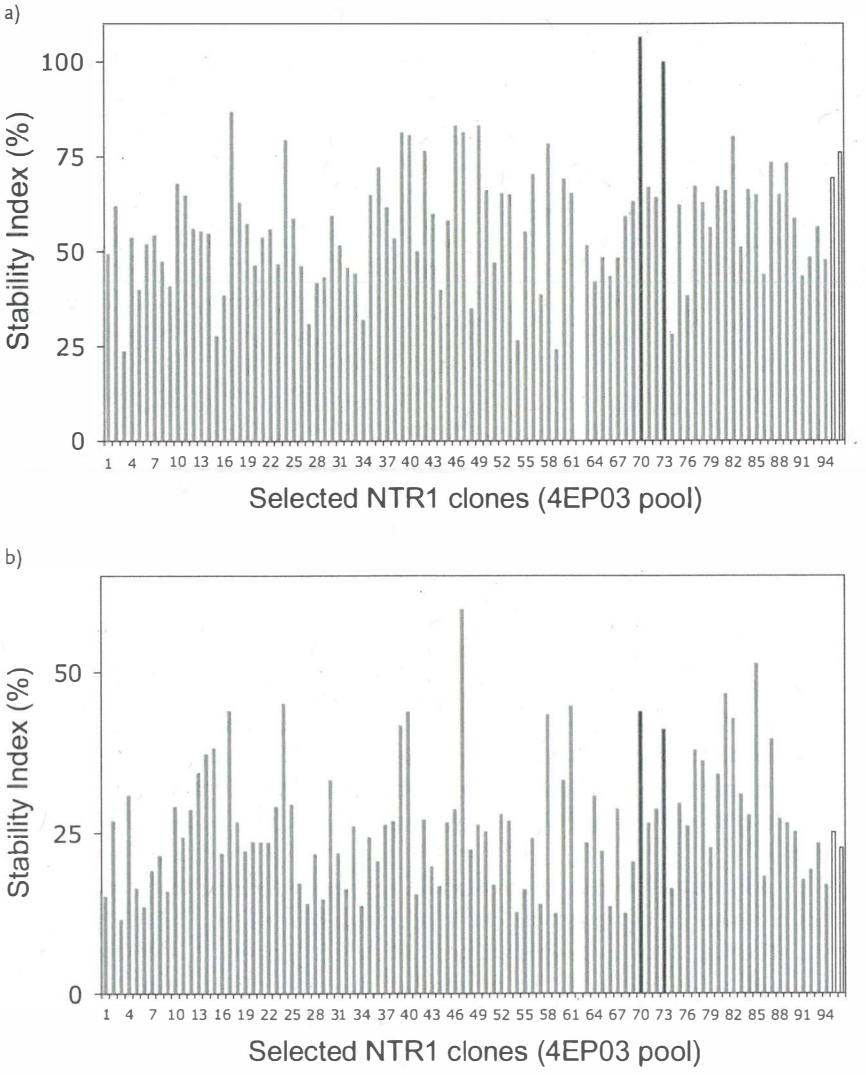

Figure 10.4 Thermal stability analysis of 96 clones isolated from selections for maximum receptor expression level (4EP03 pool, four rounds of epPCR, each followed by three to four rounds of sorting by FACS). (a and b) To identify the most thermostable clones, the stability index was measured at 37°C in (a) DDM, CHAPS, and CHS or (b) DM, CHAPS, and CHS. The stability index indicates the fraction of receptors that retain the ability to bind ligand after exposure at 37°C for 20 min (relative to receptors which are kept on ice). The open bars at positions 95 and 96 indicate duplicate measurements of the stability index of wild-type NTR1. The two black bars indicate the stability index of clone 70 and clone 73, which have been characterized in more detail. (c) The results from the screen were confirmed by reanalyzing the stability of clone 70 and clone 73 in more detail. Receptor aliquots (in DM, CHAPS, and CHS) were exposed to increasing temperatures and the fraction of intact receptors, which still binds radioligand, was measured. The resulting stability curves show that the evolved clone 70 (open triangles) and clone 73 (open circles) show a T_m (temperature at which 50% of the receptors retain ligand binding) that is 6°C higher than for wild-type NTR1 (filled circles).

c)

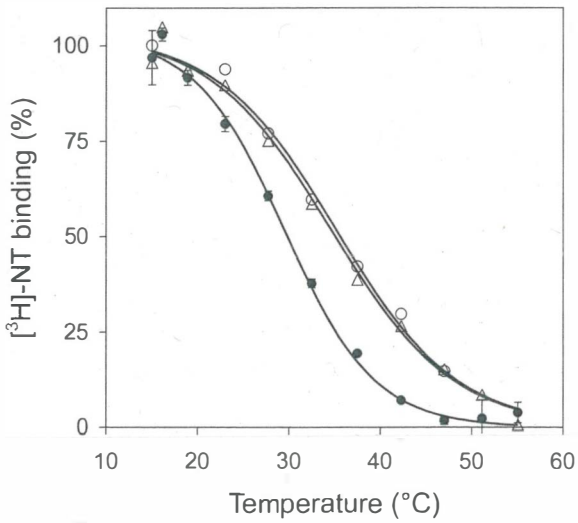

Figure 10.4 (*Continued*)

of selected mutants was strongly evolved for higher expression and not stability *per se*. In the meantime, however, from additional rounds of random mutagenesis and selection for functional expression, much higher stability increases have been found than described above or elsewhere [15] (Schlinkmann *et al.*, unpublished), suggesting that functional expression levels of GPCRs and their stability are coupled.

The principal advantage of this stability screening method based on immobilized receptors on paramagnetic beads is its high throughput. This allows for the parallelization of the rate-limiting steps in stability testing in a 96-well format, such as the LBA. Moreover, the immobilized receptors can easily be purified and exchanging detergents is straightforward, when mutants need to be screened in different detergents.

10.4 Conclusions

The presented methods for engineering the properties of IMPs all follow the same idea, namely that small changes to the amino acid sequence of a difficult-to-handle IMP can have major favorable effects. Functional expression as well as thermal stability can be increased by relatively few mutations, and can be screened and selected by comparatively simple methods. The ability to evolve functional expression level for IMPs that are very difficult to express removes one of the biggest roadblocks in IMP structural studies. Similarly, the possibility to generate IMP

variants that are more stable in a variety of detergents will be important for the future structural biology of IMPs. Two results give reason for optimism: (i) stabilizing mutations do not seem to be rare, and (ii) functional expression and stability seem to be correlated to some extent. Nonetheless, the long-term goal from applying these methods is not only to use these methods for solving the practical problem of generating IMPs that can be studied structurally more efficiently, but also apply them to understand the rules that make membrane proteins well expressed and more stable.

Abbreviations

BODIPY-NT	BODIPY-neurotensin
CHAPS	3-(3-cholamidopropyl)-dimethylammoniopropane sulfonate
CHS	cholesteryl hemisuccinate
CoFi	colony filtration
DDM	dodecyl maltoside
DGK	diacylglycerol kinase
DM	decyl maltoside
epPCR	error-prone PCR
FACS	florescence activated cell sorting
GPCR	G-protein-coupled receptor
IMP	integral membrane protein
LBA	radioligand binding assay
MBP	maltose-binding protein
MFI	mean fluorescence intensity
TrxA	thioredoxin A

References

1 Stemmer, W.P. (1994) Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature*, **370**, 389–391.

2 Zaccolo, M., Williams, D.M., Brown, D.M., and Gherardi, E. (1996) An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol.*, **255**, 589–603.

3 Virnekäs, B., Ge, L., Plückthun, A., Schneider, K.C., Wellnhofer, G., and Moroney, S.E. (1994) Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis. *Nucleic Acids Res.*, **22**, 5600–5607.

4 Molina, D.M., Cornvik, T., Eshaghi, S., Haeggstrom, J.Z., Nordlund, P., and Sabet, M.I. (2008) Engineering membrane protein overproduction in *Escherichia coli*. *Protein Sci.*, **17**, 673–680.

5 Sarkar, C.A., Dodevski, I., Kenig, M., Dudli, S., Mohr, A., Hermans, E., and Plückthun, A. (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc. Natl. Acad. Sci. USA*, **105**, 14808–14813.

6 Tucker, J. and Grisshammer, R. (1996) Purification of a rat neurotensin receptor expressed in *Escherichia coli*. *Biochem. J.*, **317**, 891–899.

7 Zhao, H., Giver, L., Shao, Z., Affholter, J.A., and Arnold, F.H. (1998) Molecular evolution by staggered extension process (StEP) *in vitro* recombination. *Nat. Biotechnol.*, **16**, 258–261.

8 Barroso, S., Richard, F., Nicolas-Etheve, D., Kitabgi, P., and Labbe-Jullie, C. (2002) Constitutive activation of the neurotensin receptor 1 by mutation of Phe^{358} in helix seven. *Br. J. Pharmacol.*, **135**, 997–1002.

9 Labbe-Jullie, C., Barroso, S., Nicolas-Eteve, D., Reversat, J.L., Botto, J.M., Mazella, J., Bernassau, J.M., and Kitabgi, P. (1998) Mutagenesis and modeling of the neurotensin receptor NTR1. Identification of residues that are critical for binding SR 48692, a nonpeptide neurotensin antagonist. *J. Biol. Chem.*, **273**, 16351–16357.

10 Middleton, R.J. and Kellam, B. (2005) Fluorophore-tagged GPCR ligands. *Curr. Opin. Chem. Biol.*, **9**, 517–525.

11 Daly, C.J. and McGrath, J.C. (2003) Fluorescent ligands, antibodies, and proteins for the study of receptors. *Pharmacol. Ther.*, **100**, 101–118.

12 Cornvik, T., Dahlroth, S.L., Magnusdottir, A., Herman, M.D., Knaust, R., Ekberg, M., and Nordlund, P. (2005) Colony filtration blot: a new screening method for soluble protein expression in *Escherichia coli*. *Nat. Methods*, **2**, 507–509.

13 Lau, F.W., Nauli, S., Zhou, Y., and Bowie, J.U. (1999) Changing single side-chains can greatly enhance the resistance of a membrane protein to irreversible inactivation. *J. Mol. Biol.*, **290**, 559–564.

14 Zhou, Y. and Bowie, J.U. (2000) Building a thermostable membrane protein. *J. Biol. Chem.*, **275**, 6975–6979.

15 Shibata, Y., White, J.F., Serrano-Vega, M.J., Magnani, F., Aloia, A.L., Grisshammer, R., and Tate, C.G. (2009) Thermostabilization of the neurotensin receptor NTS1. *J. Mol. Biol.*, **390**, 262–277.

16 Serrano-Vega, M.J., Magnani, F., Shibata, Y., and Tate, C.G. (2008) Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proc. Natl. Acad. Sci. USA*, **105**, 877–882.

17 Magnani, F., Shibata, Y., Serrano-Vega, M.J., and Tate, C.G. (2008) Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor. *Proc. Natl. Acad. Sci. USA*, **105**, 10744–10749.

18 Warne, T., Serrano-Vega, M.J., Baker, J.G., Moukhametzianov, R., Edwards, P.C., Henderson, R., Leslie, A.G., Tate, C.G., and Schertler, G.F. (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature*, **454**, 486–491.

19 Dodevski, I. and Plückthun, A. (2011) Evolution of three human GPCRs for higher expression and stability. *J. Mol. Biol.*, in press.